Using BankIt To Submit 16S rRNAs To GenBank

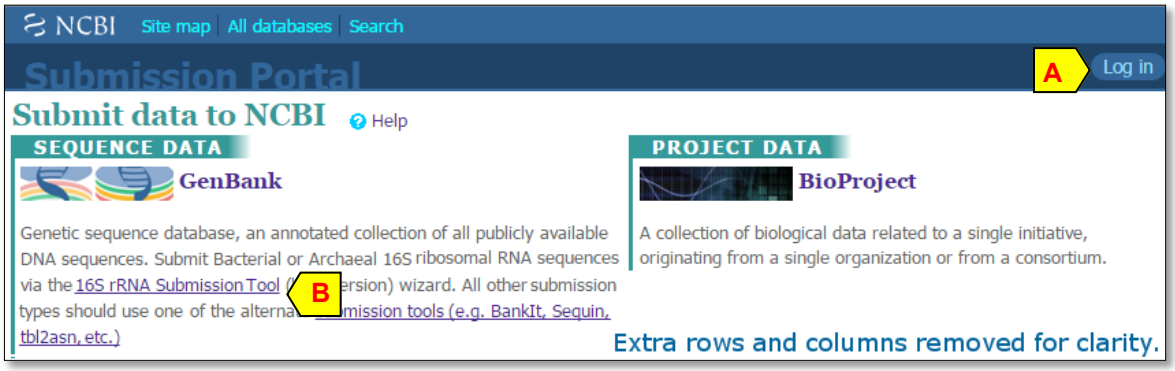
https://submit.ncbi.nlm.nih.gov/subs/genbank/

Submitting microbial 16S RNA sequences using a set of streamlined BankIt web forms

National Center for Biotechnology Information • National Library of Medicine • National Institutes of Health • Department of Health and Human Services

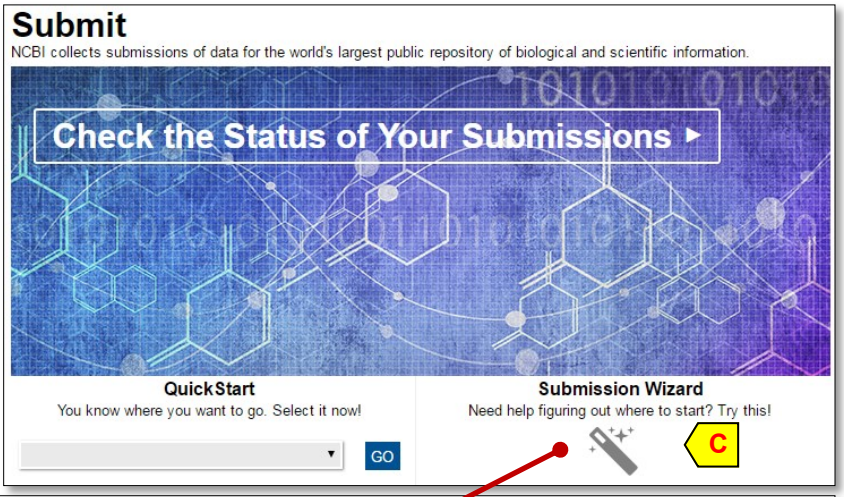
Introduction

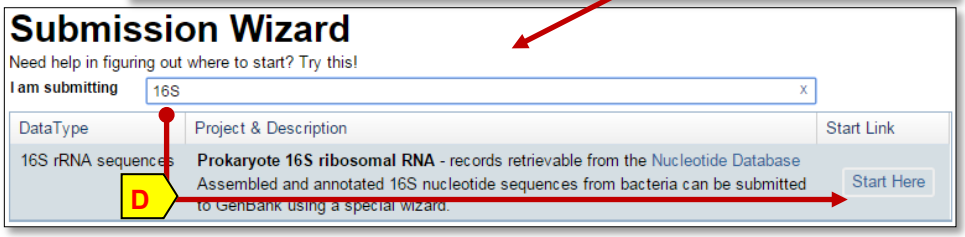
Bacterial and archaeal 16S ribosomal RNA (rRNA) sequences are extremely useful reagents for studying the composition of biologically important microbiome communities. The high degree of conservation in key regions of 16S sequences allows amplification with universal PCR primers, while the sequence diversity in other regions allows sample differentiation. A significant fraction of sequences deposited in GenBank are 16S seqiemces. BLAST searching against this collection can identify the source organisms of the input 16S rRNAs and the taxonomic composition present in the source sample.



Access

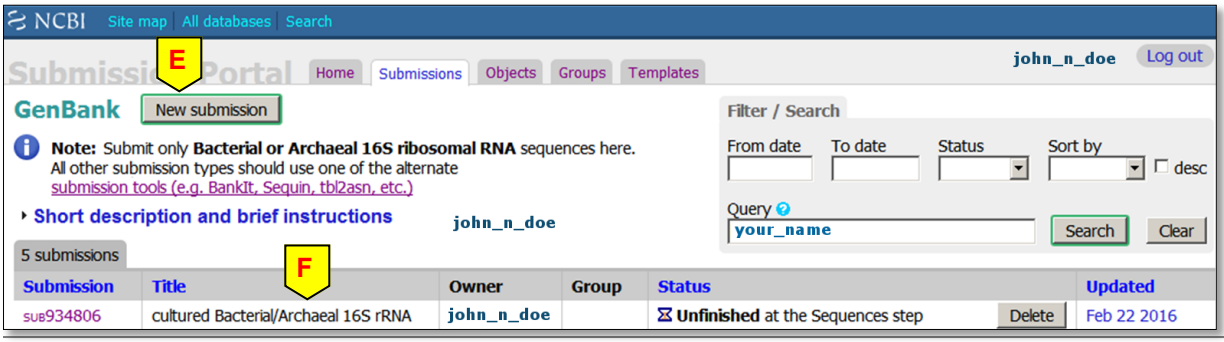
BankIt is a browser form-based tool for submitting relatively small numbers of sequences to GenBank. NCBI has streamlined the BankIt tool to for the submission of 16S rRNAs, reducing the submission to a 9-step process. BankIt is available from the Submission portal (https://submit.ncbi.nlm.nih.gov, partially shown above). You need log in to your MyNCBI account (**A**) and click the "16S rRNA Submission Tool" like (**B**) to submit your 16S rRNA sequences. Alternatively, you can also use the submission wizard (**C**) from the Submit page to locate BankIt and other submission tools. Type the data type (for example, 16S) in the text box to see suggested tools (**D**).





To start a new submission:

- Collect the sequences you want to submit in a plain text file, with sequences in the FASTA format
- Click the "16S rRNA submission Tool" link (**B**) or the "Start Here" button (**D**)
- Click the "New submission" button (**E**) in that page to begin a new submission

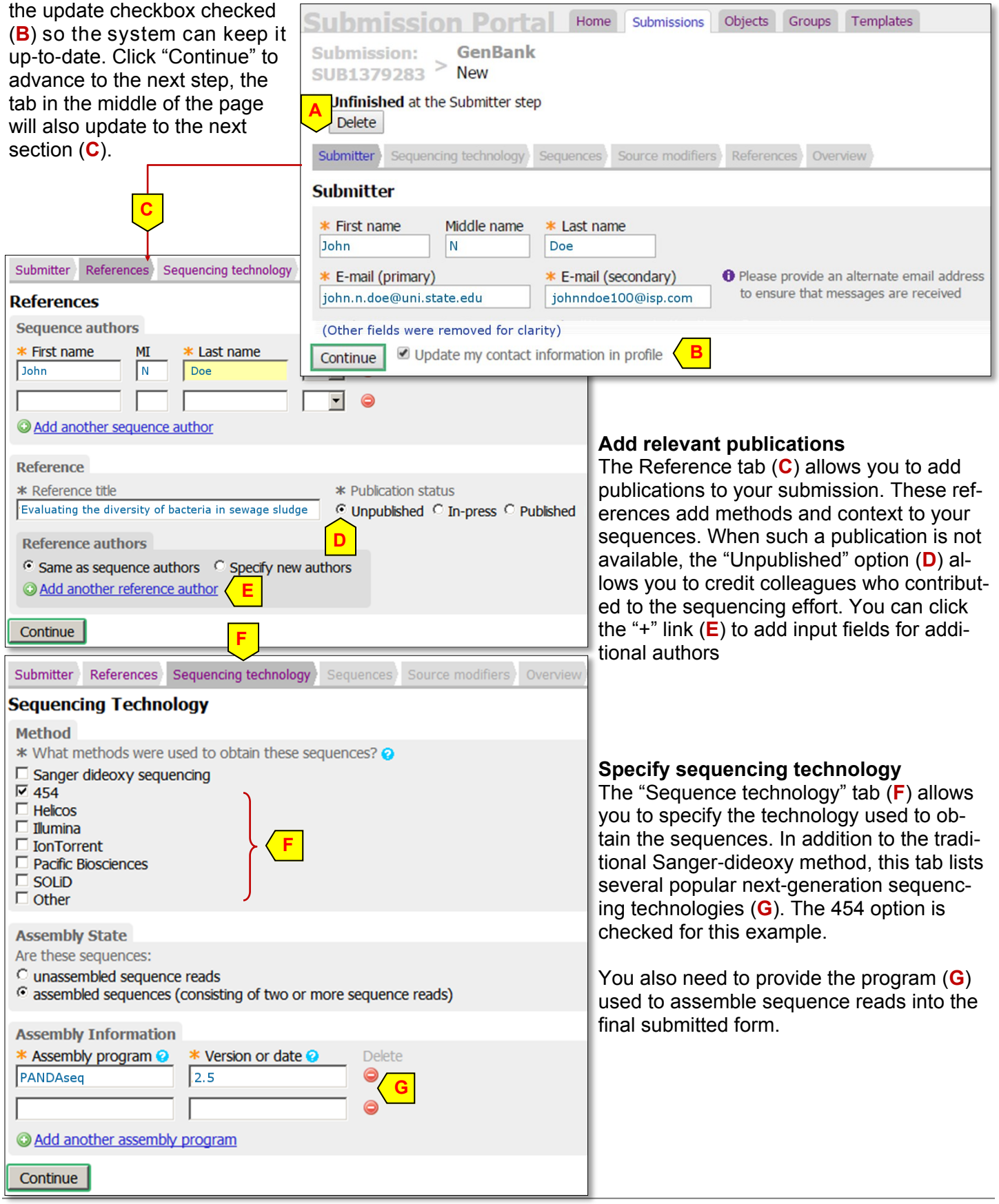


The table (**F**) at the end of the web form lists your existing submissions, their status, and date of last update.

Steps Needed To Complete A Submission

Provide your profile

After you click the "New Submission" button, the first screen you need to fill out is the "Submitter" tab (**A**, shown in part). The server will automatically populate the fields from your existing profile. Make sure you keep the update checkbox checked (**B**) so the system can keep it up-to-date. Click "Continue" to advance to the next step, the tab in the middle of the page will also update to the next section (**C**).



Add relevant publications

The Reference tab (**C**) allows you to add publications to your submission. These references add methods and context to your sequences. When such a publication is not available, the "Unpublished" option (**D**) allows you to credit colleagues who contributed to the sequencing effort. You can click the "+" link (**E**) to add input fields for additional authors

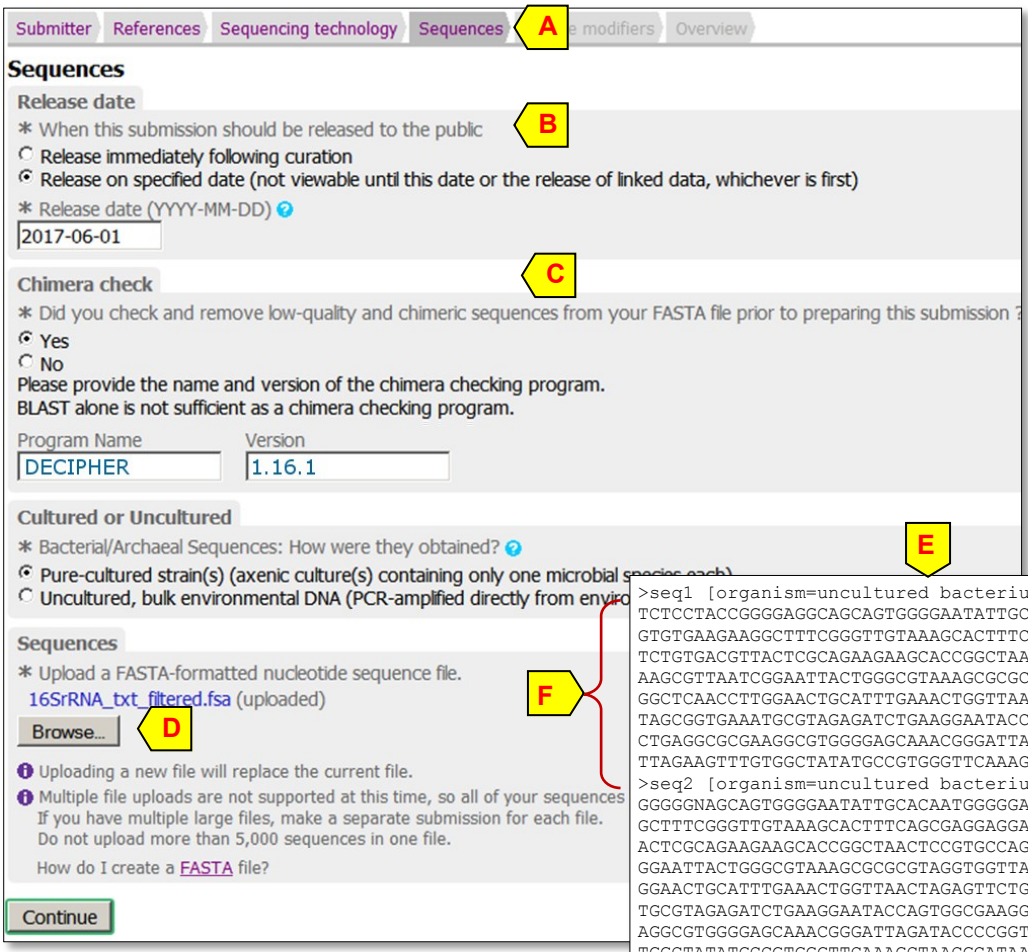
Specify sequencing technology

The "Sequence technology" tab (**F**) allows you to specify the technology used to obtain the sequences. In addition to the traditional Sanger-dideoxy method, this tab lists several popular next-generation sequencing technologies (**G**). The 454 option is checked for this example.

You also need to provide the program (**G**) used to assemble sequence reads into the final submitted form.

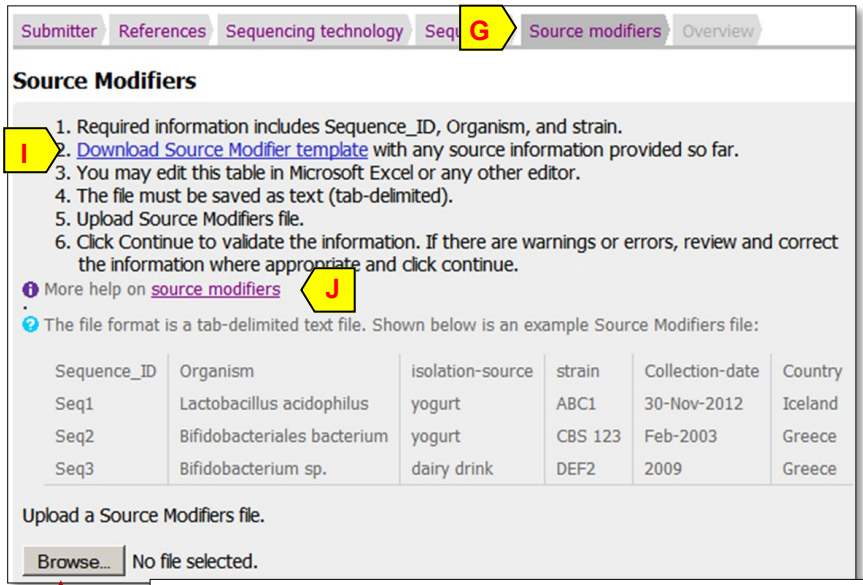
Steps Needed To Complete A Submission (cont.)

Uploading your sequences



The Sequences tab (**A**) allows you to provide more information about your sequences and to upload them. You will need to specify the public release date (**B**) for the sequences you are submitting. You should also provide details on how you checked for sequence chimeras (**C**) to ensure the accuracy of these submitted sequences. Then, upload the sequence file (**D**). The example sequences in FASTA format (**E**) contain organism names in square brackets (**F**).

Note that the 16S rRNA BankIt has a limit of a single file with up to 5000 sequences. You will need to break up larger number of sequences into smaller files and submit each separately.



Add source modifier

Feature annotation provides additional metadata adding great value to submitted sequences. For 16S rRNA sequences the most important metadata are the isolation source and taxonomic identification.
The "Source modifier" tab (**G**) allows you to provide biological source and taxonomic information to sequences you are submitting.

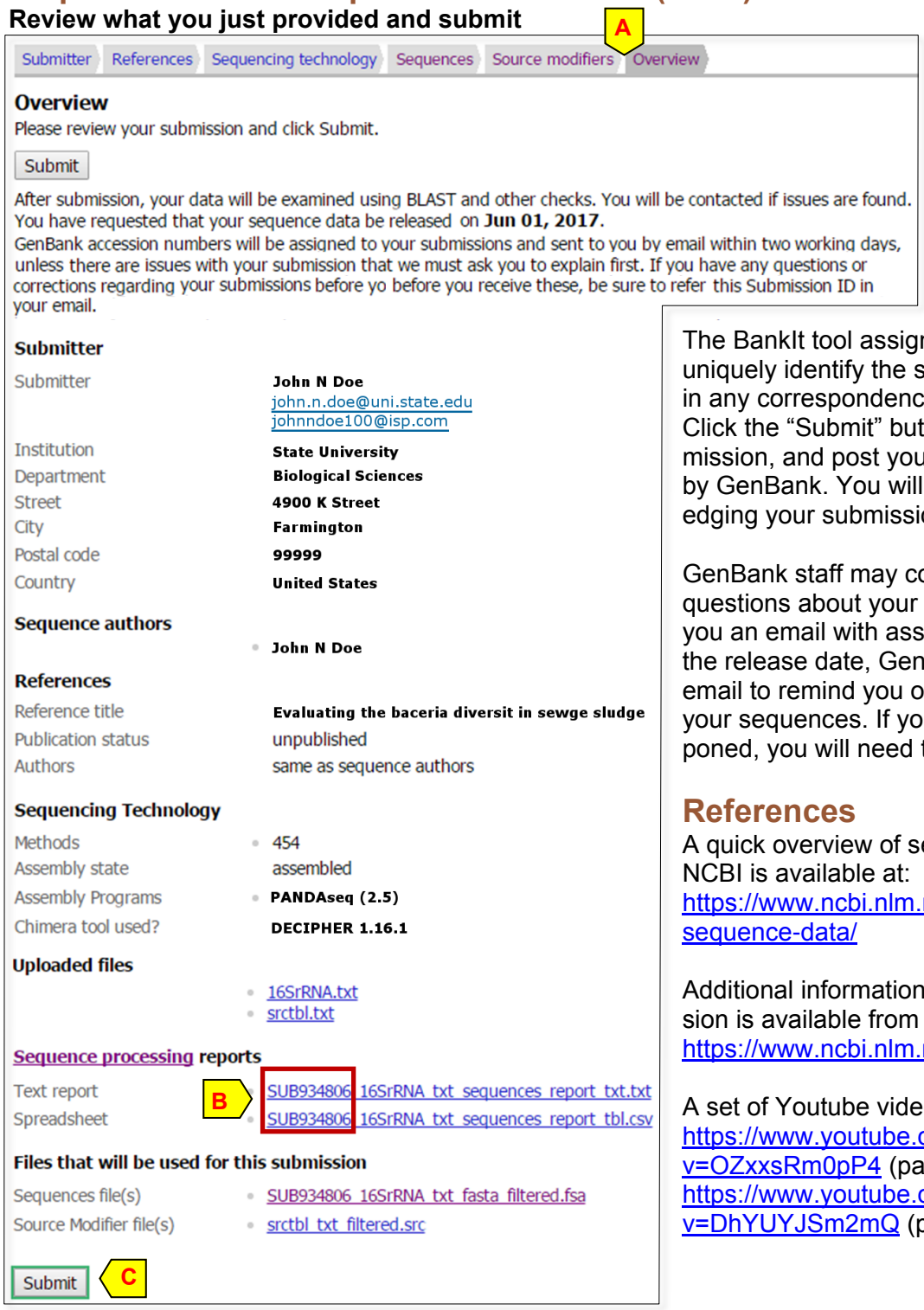
You need to upload a metadata file for your sequences in the form of tab-delimited source modifier file through the "Browse" button (**H**). You can download a template (**I**) to model your input. See the help page (**J**) for more details.

H

Sequence_ID	Isolation_source	Clone	Country	Host	Collection_date	Lat_Lon
seq1	host hemolymph	1611_16_117	Germany	Crassostrea gigas	21-May-2013	55.0289 N 8.4342 E
seq2	host hemolymph	1611_6_32	Germany	Crassostrea gigas	21-May-2013	55.0289 N 8.4342 E

Steps Needed To Complete A Submission (cont.)

Review what you just provided and submit



After you have provided all the necessary information and uplodated both the sequence and source modifier files, BankIt displays them in a single page for you to review for accuracy (**A**).

The BankIt tool assigns a submission ID (**B**) to uniquely identify the submission. Refer to this ID in any correspondence about the submission. Click the "Submit" button (**C**) to complete the submission, and post your sequences for processing by GenBank. You will receive an email acknowledging your submission.

GenBank staff may contact you if they have any questions about your submission. They will send you an email with assigned accessions. Before the release date, GenBank staff will send you an email to remind you of the pending release of your sequences. If you want the release postponed, you will need to reply and set a new date.

References

A quick overview of sequence submission to NCBI is available at:
https://www.ncbi.nlm.nih.gov/guide/howto/submit-sequence-data/

Additional information on GenBank and submission is available from the GenBank site at:
https://www.ncbi.nlm.nih.gov/genbank/

A set of Youtube video tutorials on BankIt is at:
https://www.youtube.com/watch?v=OZxxsRm0pP4 (part 1)
https://www.youtube.com/watch?v=DhYUYJSm2mQ (part 2)

Technical Support

You can send technical questions about how to use BankIt or problems encountered to:
info@ncbi.nlm.nih.gov

For questions about a successfully completed submission, you should send them to:
gb-admin@ncbi.nlm.nih.gov
Make sure you include the submission IDs assigned to your submission.